ON DIRECTION FINDING WITH UNKNOWN NOISE COVARIANCE

Benjamin Friedlander
Dept. Elec. & Comp. Eng.
University of California
Davis, CA 95616

Anthony J. Weiss
Electrical Engineering Dept.
Tel-Aviv University
Tel-Aviv, 69978, Israel

ABSTRACT

A class of direction finding methods which operate in the presence of correlated noise with an unknown covariance matrix is presented. The approach is based on joint estimation of the directions of arrival and the parameters of a model for the noise covariance matrix, using a maximum likelihood estimator, its suboptimal version or other methods. Formulas for evaluating the maximal number of identifiable noise parameters are also derived. Using the Cramer Rao bound we study the degradation in DOA estimation accuracy due to the estimation of the noise parameters.

1. INTRODUCTION

Array signal processing is used in diverse areas such as radar, sonar, communications, and seismic exploration. Usually the parameters of interest are the direction of arrival (DOA) of the observed signals and the signal waveforms. Many high resolution methods for estimating these parameters have been proposed and analyzed. Most of these methods are variations of the MUSIC algorithm [2], which requires that the noise covariance matrix be known up to a multiplicative factor. In many radio frequency (RF) systems the dominant noise is the thermal noise which is approximately equal in all the channels. In these cases the correct noise covariance is a scaled identity matrix (or a diagonal matrix, if the channels are unequal), and the assumption of known noise covariance is justified.

In this paper we focus on the case where the dominant noise is external (ambient) noise. Ambient noise is the dominant noise source in RF systems operating in the HF and VHF frequency bands, and in most sonar systems [1], and its presence introduces correlation between the noise processes of the different sensors. This has led to the development of methods which attempt to take into account the presence of unknown noise correlation.

One such method is the covariance differencing technique proposed in [3]. In this method two measurements of the array covariance are required, where the unknown noise covariance remains invariant, while the signal component undergoes some change between the two measurements. Then, the unknown noise covariance is eliminated by subtraction. However, as stated in [8], in many practical situations, only one measurement of the covariance is available.

Recently, a MAP approach was presented in [8, 9]. This method assumes that the noise covariance matrix is completely unknown, except for the fact that it is a Hermitian positive definite matrix. This method is reported to perform better than MUSIC and Maximum Likelihood which are applied using the (wrong) assumption that the noise covariance is a scaled identity matrix. In [9] it is shown that the MAP estimator proposed in [8] is in general inconsistent, except in special cases. A similar formulation that assumes completely unknown noise covariance was used in [7]. Wax advocated a technique based on Rissanen's MDL principle for detection and localization of signals in noise with completely unknown noise covariance. However, this technique is asymptotically biased, as is the case in [8].

The authors of [6] and [10] proposed a parameterization of the noise covariance based on an AR model. The method proposed in [6] and the technique suggested in [10] for DOA estimation are both limited to linear uniform arrays. A formulation that is close in spirit to the formulation developed in this work is presented in [4]. While [4] concentrates on numerical techniques for DOA estimation, we focus on the Cramer Rao bound and the best performance that can be achieved.

In this work we show that in the presence of ambient noise, the noise covariance matrix can be represented as sum of known matrices which are dependent on the array configuration. These matrices are multiplied by a an unknown constant whose size is determined by the intensity and the spatial distribution of the noise. Using the Cramer Rao bound, we examine the number of unknown noise parameters that can be estimated, the accuracy of the DOA estimates as a function of the number of noise parameters, and the effect of unknown noise covariance on the DOA estimates. In [11] we evaluated by simulations the

This work was supported by the Army Research Office, contract no. DAAL03-91-C-0022, sponsored by U.S. Army Communications Electronics Command, Center for Signals Warfare.

1058-6393/95 $4.00 © 1995 IEEE

performance of a Maximum Likelihood estimator, and showed that its performance is very close to the CRB.

2. PROBLEM FORMULATION

We describe the data model for the narrowband DOA estimation problem. To simplify the exposition our discussion is confined to azimuth-only systems, *i.e.*, the sensors and signals are assumed to be coplanar.

We consider an M-element array of sensors and L narrowband far-field signal sources, and define the $M \times 1$ vector $\mathbf{a}(\theta)$ to be the complex array response for a source at direction θ. The array manifold is defined to be the continuum $\mathcal{M} = \{\mathbf{a}(\theta) : \theta \in [-\pi, \pi]\}$.

The outputs of the M array elements at the k-th sample are arranged in an $M \times 1$ vector,

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{n}(k) \qquad k = 1, 2 \cdots N; \tag{1}$$

where $\mathbf{n}(k)$ is the noise vector, $\mathbf{s}(k)$ is the signal vector, and

$$\mathbf{A} \triangleq [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \cdots, \mathbf{a}(\theta_L)] \tag{2}$$

Assuming that the signal vectors and the noise vectors are realizations of stationary, zero mean Gaussian random processes, and that the noise and the signals are uncorrelated, the data covariance matrix is

$$\mathbf{R} \triangleq \mathrm{E}\{\mathbf{x}(k)\mathbf{x}^H(k)\} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{\Sigma} \tag{3}$$

where $\mathbf{P}$ is the signal covariance matrix and $\mathbf{\Sigma}$ is the noise covariance matrix.

Modeling of the Noise Covariance Matrix

The noise in most receiving systems consists of internal noise and external noise. The internal noise is produced by the electronic equipment and includes thermal noise and weak versions of other signals in the system like clocks and local oscillators. The external noise is defined as an unwanted random signal that is intercepted by the sensor. If the system is designed well, so that there is no coupling between the receiving sensors, and the thermal noise is the main source of internal noise, then a good model for the internal noise covariance is a scaled identity matrix, $\mu\mathbf{I}$. This model assumes that the thermal noise intensity is the same in all sensors and that there is no correlation between the noise at any two sensors.

Next we define an appropriate model for the external noise. The noise intensity as a function of azimuth θ at a given time instant k is a random function denoted by $v(\theta|k)$. Therefore, the noise vector at the array output (assuming external noise only) is given by

$$\mathbf{n}(k) = \int_{-\pi}^{\pi} \mathbf{a}(\theta) v(\theta|k)\, d\theta \tag{4}$$

We assume that $v(\theta|k)$ satisfies the following relations

$$\mathrm{E}\{v(\theta|k)v^*(\sigma|j)\} = \varepsilon(\theta)\delta(\theta - \sigma)\delta_{kj} \tag{5}$$

$$\mathrm{E}\{v(\theta|k)v(\sigma|j)\} = 0 \tag{6}$$

where δ_{kj} is the Kronecker delta, $\delta(\tau)$ is Dirac's delta and $\varepsilon(\theta)$ is the spatial power density function of the noise. Using these relations we get for the external noise covariance

$$\begin{aligned}
\mathrm{E}\{\mathbf{n}(k)\mathbf{n}^H(k)\} = &\\
&\mathrm{E}\left\{\int_{-\pi}^{\pi} \mathbf{a}(\theta)v(\theta|k)\, d\theta \int_{-\pi}^{\pi} \mathbf{a}(\sigma)v(\sigma|k)\, d\sigma\right\}\\
&= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathbf{a}(\theta)\mathbf{a}^H(\sigma)\varepsilon(\theta)\delta(\theta-\sigma)\, d\theta d\sigma\\
&= \int_{-\pi}^{\pi} \mathbf{a}(\theta)\mathbf{a}^H(\theta)\varepsilon(\theta)\, d\theta
\end{aligned} \tag{7}$$

Since $\varepsilon(\theta)$ is a periodic function it can be represented by Fourier series as follows,

$$\varepsilon(\theta) = \sum_{\ell=0}^{\infty} c_\ell \cos(\ell\theta) + s_\ell \sin(\ell\theta) \tag{8}$$

where

$$c_\ell \triangleq \begin{cases} \dfrac{1}{2\pi}\displaystyle\int_{2\pi} \varepsilon(\theta)\, d\theta & \text{for } \ell = 0 \\ \dfrac{1}{\pi}\displaystyle\int_{2\pi} \varepsilon(\theta)\cos(\ell\theta)\, d\theta & \text{for all } \ell > 0 \end{cases} \tag{9}$$

$$s_\ell \triangleq \begin{cases} 0 & \text{for } \ell = 0 \\ \dfrac{1}{\pi}\displaystyle\int_{2\pi} \varepsilon(\theta)\sin(\ell\theta)\, d\theta & \text{for all } \ell > 0 \end{cases} \tag{10}$$

In general the Fourier series representation of $\varepsilon(\theta)$ will have an infinite number of terms. However, in most practical cases of interest the noise power spectral density function is smooth and varies slowly with direction. Thus, $\varepsilon(\theta)$ can be usually approximated quite well by a Fourier series with a small number of terms, say, $\bar{L}$.

Substituting (8) in (7) and assuming that for $\ell > \bar{L}$ the coefficients of the Fourier series are zero, we get,

$$\mathrm{E}\{\mathbf{n}(k)\mathbf{n}^H(k)\} = \sum_{\ell=0}^{\bar{L}} c_\ell \bar{\mathbf{\Sigma}}_\ell + s_\ell \tilde{\mathbf{\Sigma}}_\ell \tag{11}$$

where

$$\bar{\mathbf{\Sigma}}_\ell \triangleq \int_{-\pi}^{\pi} \mathbf{a}(\theta)\mathbf{a}^H(\theta)\cos(\ell\theta)\, d\theta \tag{12}$$

$$\tilde{\mathbf{\Sigma}}_\ell \triangleq \int_{-\pi}^{\pi} \mathbf{a}(\theta)\mathbf{a}^H(\theta)\sin(\ell\theta)\, d\theta \tag{13}$$

All $\bar{\Sigma}_\ell$ and $\tilde{\Sigma}_\ell$ are Hermitian. It can be shown, that for any 2D array, with omnidirectional sensors, the following relations hold,

$$\mathrm{Im}\{\bar{\Sigma}_\ell\} = \mathrm{Im}\{\tilde{\Sigma}_\ell\} = 0, \qquad \ell = 2n, \tag{14}$$

$$\mathrm{Re}\{\bar{\Sigma}_\ell\} = \mathrm{Re}\{\tilde{\Sigma}_\ell\} = 0, \qquad \ell = 2n+1, \tag{15}$$

where $n = 0, 1, 2, \cdots$ Also, for any 1D array (linear array), with omnidirectional sensors, the following relations hold,

$$\mathrm{Im}\{\bar{\Sigma}_\ell\} = 0, \qquad \ell = 2n+1, \tag{16}$$

$$\mathrm{Re}\{\tilde{\Sigma}_\ell\} = 0, \qquad \ell = 2n, \tag{17}$$

We therefore conclude that for linear array, with omnidirectional sensors, the matrices $\bar{\Sigma}_\ell$ for odd ℓ, and the matrices $\tilde{\Sigma}_\ell$ for even ℓ, vanish.

Note that it is impossible to uniquely identify the coefficients c_ℓ and s_ℓ if the noise matrices are not linearly independent. In general, the noise matrices are confined to a M^2-dimensional real linear space due to the fact that these matrices are Hermitian. However, for certain arrays this space may be considerably smaller. Consider for example a uniform linear array consisting of omnidirectional sensors. In this case the matrices $\bar{\Sigma}_\ell$, $\tilde{\Sigma}_\ell$ are all Toeplitz, and therefore they are confined to a $(2M-1)$-dimensional real space. Thus, in the general case we can not estimate more than M^2 noise parameters and in the case of uniform linear array with omnidirectional sensors we are limited to $2M-1$ noise parameters. Further restrictions on the number of noise parameters are presented later.

Our basic assumption in this work is that both the noise and the signals are realizations of Gaussian processes. Therefore, all the information about the parameters of interest is embedded in the data covariance matrix $\mathbf{R}$. Intuitively, any estimation scheme should look for parameter values that "best" fit the model of $\mathbf{R}$. The number of independent parameters that define $\mathbf{R}$ is M^2 since the covariance matrix is Hermitian. Therefore, intuitively, the maximum parameters that can be estimated is M^2. On the other side, the number of parameters that should be estimated is as follows. The L directions of arrival, the L^2 independent parameters of $\mathbf{P}$, and the $\tilde{J}$ noise parameters. Thus, the problem is well posed only if

$$M^2 \geq L + L^2 + \tilde{J} \tag{18}$$

Therefore, we can not expect to estimate the directions of arrival properly, if there are more than $M^2 - L^2 - L$ noise parameters. This intuitive observation is supported by the Cramer Rao bound. The Fisher Information Matrix become singular when the relation (18) is violated. A full derivation of the Cramer Rao bound can be found in [11].

3. MAXIMUM LIKELIHOOD ESTIMATION

Once we have a parametric model for the problem we can apply various estimation techniques. Each of these algorithms has different advantages and limitations compared with the others. In this section we derive two such algorithms: an exact maximum likelihood algorithm, and a suboptimal but computationally more efficient version.

It is well known that the Maximum Likelihood estimator is asymptotically (with large number of samples) unbiased and statistically efficient. Since we assumed that the signals and the noise processes are Gaussian the probability density function of the data is given by

$$f(\mathbf{x}(k)|\boldsymbol{\phi}) = \frac{1}{\det\{\pi\mathbf{R}\}} \exp\{-\mathbf{x}(k)\mathbf{R}^{-1}\mathbf{x}(k)\} \tag{19}$$

where $\boldsymbol{\phi}$ is the parameter vector. In our case $\boldsymbol{\phi}$ includes the directions of arrival, the entries of $\mathbf{P}$ and the noise parameters $\{\eta_j\}$. If we collect N independent data samples the log likelihood function is

$$Q(\boldsymbol{\phi}) = -N \log\{\det\{\pi\mathbf{R}\}\} - N\mathrm{tr}\{\mathbf{R}^{-1}\mathbf{D}\} \tag{20}$$

where

$$\mathbf{D} \triangleq \frac{1}{N}\sum_{k=1}^{N} \mathbf{x}(k)\mathbf{x}^H(k) \tag{21}$$

The maximizer of $Q(\boldsymbol{\phi})$ is the maximum likelihood estimator for the problem at hand. One can try to maximize (20) directly, or find ways to reduce the dimension of the multidimensional search implied by (20). In general, The dimension of the search is $L + L^2 + \tilde{J}$. However, we can reduce the search to $L + \tilde{J}$ by solving for $\mathbf{P}$ in terms of $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ and then substituting back to $Q(\boldsymbol{\phi})$. The derivation follows the steps in [5].

Maximization of (20) with respect to $\boldsymbol{\theta}, \boldsymbol{\eta}$ and $\mathbf{P}$ is equivalent to the following two-step procedure. We first maximize the function with respect to $\mathbf{P}$ for a fixed $\boldsymbol{\theta}, \boldsymbol{\eta}$ and substitute for the resulting $\hat{\mathbf{P}}$ as a function of $\boldsymbol{\theta}, \boldsymbol{\eta}$ and the data $\mathbf{D}$ back into the log-likelihood function, resulting in a function to be maximized over only the $\boldsymbol{\theta}, \boldsymbol{\eta}$ parameters. A necessary condition for an extremum of $Q(\boldsymbol{\phi})$ with respect to $\mathbf{P}$ is that the partial derivatives of $Q(\boldsymbol{\phi})$ with respect to the entries of $\mathbf{P}$ be equated to zero. Jaffer [5] shows that this condition is equivalent to

$$\mathbf{A}^H\mathbf{R}^{-1}[\mathbf{D} - \mathbf{R}]\mathbf{R}^{-1}\mathbf{A} = 0 \tag{22}$$

In order to proceed we need an expression for the inverse of $\mathbf{R}$. Define

$$\bar{\mathbf{R}} \triangleq \mathbf{\Sigma}^{-1/2}\mathbf{R}\mathbf{\Sigma}^{-1/2} = \bar{\mathbf{A}}\mathbf{P}\bar{\mathbf{A}}^H + \mathbf{I} \quad (23)$$

$$\bar{\mathbf{A}} \triangleq \mathbf{\Sigma}^{-1/2}\mathbf{A} \quad (24)$$

Using a well know identity, we get

$$\bar{\mathbf{R}}^{-1} = \mathbf{I} - \bar{\mathbf{A}}(\mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I})^{-1}\mathbf{P}\bar{\mathbf{A}}^H \quad (25)$$

and

$$\mathbf{R}^{-1} = \mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1/2}\bar{\mathbf{A}}(\mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I})^{-1}\mathbf{P}\bar{\mathbf{A}}^H\mathbf{\Sigma}^{-1/2} \quad (26)$$

which yield

$$\begin{aligned}\mathbf{R}^{-1}\mathbf{A} &= \mathbf{\Sigma}^{-1/2}\bar{\mathbf{A}}[\mathbf{I} - (\mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I})^{-1}\mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}}] \\ &= \mathbf{\Sigma}^{-1/2}\bar{\mathbf{A}}(\mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I})^{-1}[(\mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I}) - \mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}}] \\ &= \mathbf{\Sigma}^{-1/2}\bar{\mathbf{A}}(\mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I})^{-1} \quad (27)\end{aligned}$$

Substituting (27) in (22) we get

$$(\mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I})^{-1}\bar{\mathbf{A}}\mathbf{\Sigma}^{-1/2}(\mathbf{D} - \mathbf{R})\cdot \mathbf{\Sigma}^{-1/2}\bar{\mathbf{A}}(\mathbf{P}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I})^{-1} = 0 \quad (28)$$

This reduces to

$$\bar{\mathbf{A}}\mathbf{\Sigma}^{-1/2}(\mathbf{D} - \mathbf{R})\mathbf{\Sigma}^{-1/2}\bar{\mathbf{A}} = 0 \quad (29)$$

Substituting for $\mathbf{R}$ from (3) we get

$$\begin{aligned}\bar{\mathbf{A}}^H\mathbf{\Sigma}^{-1/2}\mathbf{D}\mathbf{\Sigma}^{-1/2}\bar{\mathbf{A}} &= \bar{\mathbf{A}}^H\mathbf{\Sigma}^{-1/2}(\mathbf{A}\mathbf{P}\mathbf{A} + \mathbf{\Sigma})\mathbf{\Sigma}^{-1/2}\bar{\mathbf{A}} \\ &= \bar{\mathbf{A}}^H\bar{\mathbf{A}}\mathbf{P}\mathbf{A}^H\bar{\mathbf{A}} + \bar{\mathbf{A}}^H\bar{\mathbf{A}} \quad (30)\end{aligned}$$

Thus the solution for $\mathbf{P}$ is given by

$$\hat{\mathbf{P}} = (\bar{\mathbf{A}}^H\bar{\mathbf{A}})^{-1}[\bar{\mathbf{A}}^H\bar{\mathbf{D}}\bar{\mathbf{A}} - \bar{\mathbf{A}}^H\bar{\mathbf{A}}](\bar{\mathbf{A}}^H\bar{\mathbf{A}})^{-1} \quad (31)$$

$$\bar{\mathbf{D}} \triangleq \mathbf{\Sigma}^{-1/2}\mathbf{D}\mathbf{\Sigma}^{-1/2} \quad (32)$$

Define

$$\hat{\mathbf{R}} \triangleq \mathbf{A}\hat{\mathbf{P}}\mathbf{A}^H + \mathbf{\Sigma} \quad (33)$$

Then the maximization of $Q(\boldsymbol{\phi})$ is reduced to the maximization of

$$Q_1(\boldsymbol{\theta}, \boldsymbol{\eta}) = -\log\{\det\{\hat{\mathbf{R}}\}\} - \mathrm{tr}\{\hat{\mathbf{R}}^{-1}\mathbf{D}\} \quad (34)$$

with respect to $\boldsymbol{\theta}, \boldsymbol{\eta}$. Unfortunately, equation (34) requires a matrix inversion. We can simplify (34) following the steps described in [5] with the required modifications. Using (26) we have

$$\hat{\mathbf{R}}^{-1}\mathbf{D} = \mathbf{\Sigma}^{-1}\mathbf{D} - \mathbf{\Sigma}^{-1/2}\bar{\mathbf{A}}(\hat{\mathbf{P}}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I})^{-1}\hat{\mathbf{P}}\bar{\mathbf{A}}^H\mathbf{\Sigma}^{-1/2}\mathbf{D} \quad (35)$$

Substituting (31) in $\hat{\mathbf{P}}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I}$ we get

$$\hat{\mathbf{P}}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I} = (\bar{\mathbf{A}}^H\bar{\mathbf{A}})^{-1}\bar{\mathbf{A}}^H\bar{\mathbf{D}}\bar{\mathbf{A}} \quad (36)$$

Taking the inverse of this result yield

$$(\hat{\mathbf{P}}\bar{\mathbf{A}}^H\bar{\mathbf{A}} + \mathbf{I})^{-1} = (\bar{\mathbf{A}}^H\bar{\mathbf{D}}\bar{\mathbf{A}})^{-1}\bar{\mathbf{A}}^H\bar{\mathbf{A}} \quad (37)$$

Substituting (37) in (35) and evaluating the trace of the matrices, we get

$$\mathrm{tr}\{\hat{\mathbf{R}}^{-1}\mathbf{D}\} = \mathrm{tr}\{\bar{\mathbf{D}}\} - \mathrm{tr}\{\bar{\mathbf{A}}^H\bar{\mathbf{A}}\hat{\mathbf{P}}\} \quad (38)$$

Using (36) one gets

$$\begin{aligned}\mathrm{tr}\{\hat{\mathbf{P}}\bar{\mathbf{A}}^H\bar{\mathbf{A}}\} &= \mathrm{tr}\{(\bar{\mathbf{A}}^H\bar{\mathbf{A}})^{-1}\bar{\mathbf{A}}^H\bar{\mathbf{D}}\bar{\mathbf{A}} - \mathbf{I}\} \\ &= \mathrm{tr}\{\mathbf{G}\bar{\mathbf{D}}\} - L \quad (39)\end{aligned}$$

where

$$\mathbf{G} \triangleq \bar{\mathbf{A}}(\bar{\mathbf{A}}^H\bar{\mathbf{A}})^{-1}\bar{\mathbf{A}}^H \quad (40)$$

Using this result equation (38) becomes

$$\mathrm{tr}\{\hat{\mathbf{R}}^{-1}\mathbf{D}\} = \mathrm{tr}\{\mathbf{H}\bar{\mathbf{D}}\} + L, \qquad \mathbf{H} \triangleq \mathbf{I} - \mathbf{G} \quad (41)$$

The matrices $\mathbf{G}$ and $\mathbf{H}$ are recognized as projection matrices on the column space of $\bar{\mathbf{A}}$ and on the null space of $\bar{\mathbf{A}}^H$, respectively.

Substituting (31) in (33) and multiplying by $\mathbf{\Sigma}^{-1/2}$, we get

$$\mathbf{\Sigma}^{-1/2}\hat{\mathbf{R}}\mathbf{\Sigma}^{-1/2} = \mathbf{G}\bar{\mathbf{D}}\mathbf{G} + \mathbf{H} \quad (42)$$

Also

$$\det\{\mathbf{\Sigma}^{-1/2}\hat{\mathbf{R}}\mathbf{\Sigma}^{-1/2}\} = \frac{\det\{\hat{\mathbf{R}}\}}{\det\{\mathbf{\Sigma}\}} \quad (43)$$

Hence, (43) and (42) yield

$$\det\{\hat{\mathbf{R}}\} = \det\{\mathbf{\Sigma}\}\det\{\mathbf{G}\bar{\mathbf{D}}\mathbf{G} + \mathbf{H}\} \quad (44)$$

Substituting (44) and (41) in (34) and eliminating the constant L that does not affect the maximization, we get

$$Q_2(\boldsymbol{\theta}, \boldsymbol{\eta}) = -\log\{\det\{\mathbf{\Sigma}\}\} - \log\{\det\{\mathbf{G}\bar{\mathbf{D}}\mathbf{G} + \mathbf{H}\}\} - \mathrm{tr}\{\mathbf{H}\bar{\mathbf{D}}\} \quad (45)$$

The maximization of Q_2 requires a reduced dimension search. We have tried to reduce the search dimension even more by expressing the noise parameters estimates using only the directions of arrival. Unfortunately, we could do this only for a single noise parameter, in which case we get

$$\hat{\eta}_0 = \frac{1}{M - L}\mathrm{tr}\{\mathbf{H}\mathbf{\Sigma}_0^{-1/2}\mathbf{S}\mathbf{\Sigma}_0^{-1/2}\} \quad (46)$$

The generalization of this result is still an open problem.

4. SUBOPTIMAL ESTIMATION

In the previous section we described the exact Maximum Likelihood (ML) algorithm for the problem at hand. The ML estimator is known to be asymptotically unbiased and efficient. Unfortunately, this method requires a search in a parameter space whose dimension is $L+\tilde{J}$. We now derive a suboptimal algorithm, using intuitive arguments, whose performance is acceptable, in most cases. The search associated with this algorithm is of dimension L.

We first note that

$$\lim_{N\to\infty} \mathbf{D} = \mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \boldsymbol{\Sigma} \tag{47}$$

Therefore, a reasonable cost function for minimization is

$$Q(\boldsymbol{\phi}) = ||\mathbf{D} - \mathbf{A}\mathbf{P}\mathbf{A}^H - \boldsymbol{\Sigma}||^2 = ||\mathbf{d} - \mathbf{B}\mathbf{g} - \boldsymbol{\Gamma}\boldsymbol{\eta}||^2 \tag{48}$$

where

$$\mathbf{d} \triangleq \text{vec}\{\mathbf{D}\} \tag{49}$$

$$\mathbf{B} \triangleq \mathbf{A}^* \otimes \mathbf{A} \tag{50}$$

$$\mathbf{g} \triangleq \text{vec}\{\mathbf{P}\} \tag{51}$$

$$\boldsymbol{\Gamma} \triangleq [\text{vec}\{\boldsymbol{\Sigma}_0\}, \text{vec}\{\boldsymbol{\Sigma}_1\}, \cdots \text{vec}\{\boldsymbol{\Sigma}_J\}] \tag{52}$$

The minimization of (48) with respect to $\boldsymbol{\eta}$ is obtained by choosing

$$\hat{\boldsymbol{\eta}} = (\boldsymbol{\Gamma}^H\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}^H(\mathbf{d} - \mathbf{B}\mathbf{g}) \tag{53}$$

Substituting (53) back to (48) and using the definition

$$\boldsymbol{\Pi} \triangleq \mathbf{I} - \boldsymbol{\Gamma}(\boldsymbol{\Gamma}^H\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}^H \tag{54}$$

we get

$$Q = ||\boldsymbol{\Pi}(\mathbf{d} - \mathbf{B}\mathbf{g}||^2 \tag{55}$$

The minimization of (55) with respect to $\mathbf{g}$ is obtained by choosing

$$\hat{\mathbf{g}} = (\mathbf{B}^H\boldsymbol{\Pi}\mathbf{B})^{-1}\mathbf{B}^H\boldsymbol{\Pi}\mathbf{d} \tag{56}$$

Substituting (56) back to (55) we get

$$Q(\boldsymbol{\theta}) = ||\boldsymbol{\Pi}\left(\mathbf{I} - \mathbf{B}(\mathbf{B}^H\boldsymbol{\Pi}\mathbf{B})^{-1}\mathbf{B}^H\boldsymbol{\Pi}\right)\mathbf{d}||^2 \tag{57}$$

The minimization of (57) requires a search of dimension L which may be considerably less than the search dimension of the ML estimator. It is clear from the method of derivation that the estimation error reduces to zero as $N \to \infty$. Therefore, our estimates are asymptotically unbiased.

REFERENCES

[1] W. S. Burdic, *Underwater Acoustic System Analysis*, Prentice Hall, 1984.

[2] R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Trans. Antennas and Propag.*, vol. AP-34, no. 3, pp. 276–280, March 1986.

[3] A. Paulraj, and T. Kailath, "Eigenstructure Methods for Direction of Arrival Estimation in the Presence of Unknown Noise Fields," *IEEE Trans. ASSP*, Vol. 34, No. 1, pp. 13-20, Jan., 1986.

[4] J.F. Bohme, and D. Kraus, "On Least Squares Methods for Direction of Arrival Estimation in the Presence of Unknown Noise Fields," *Proceedings of ICASSP*, New-York, NY, pp. 2833-2836, April 1988.

[5] A.G. Jaffer, "Maximum Likelihood Direction Finding of Stochastic Sources: A Separable Solution," *Proceedings of ICASSP*, New-York, NY, pp. 2893-2896, April 1988.

[6] A. H. Tewfik, "Direction Finding in the presence of Colored Noise by Candidate Identification," *IEEE Trans. on Signal Processing*, vol. 39, no. 9, pp. 1933–1942, September 1991.

[7] M. Wax, "Detection and Localization of Multiple Sources in Noise with Unknown Covariance," *IEEE Trans. on Signal Processing*, vol. 40, no. 1, pp. 245–249, September 1991.

[8] K.M. Wong, et. al., "Estimation of the Directions of Arrival of Signals in Unknown Correlated Noise, Part I: The MAP Approach and its Implementation," *IEEE Trans. Signal Processing*, Vol. 40, No. 8, pp. 2007-2017, Aug., 1992.

[9] J.P. Reilly, and K.M. Wong, "Estimation of the Directions of Arrival of Signals in Unknown Correlated Noise, Part II: Asymptotic Behavior And Performance of the MAP Approach," *IEEE Trans. Signal Processing*, Vol. 40, No. 8, pp. 2018-2028, Aug., 1992.

[10] V. Nagesha, and S. Kay, "Maximum Likelihood Estimation for Array Processing in Colored Noise," *Proceedings of ICASSP*, Minneapolis, MN, pp. IV-240-243, April 1993.

[11] B. Friedlander, and A. J. Weiss, "Direction Finding Using Noise Covariance Modeling," submitted for publication, 1993.